American Journal of Epidemiology, 1998;148:1-3

The Human Genome Epidemiology Network (HuGENet™)

by Muin J. Khoury M.D., Ph.D. and Janice S. Dorman

The Human Genome Project and the need for population-based epidemiologic research

Progress of the Human Genome Project has led to an explosion of genetic information (1). Of the estimated 100,000 human genes, more than 9,000 have been discovered, and more than 5,000 have been mapped to specific chromosomes (2). There are about 500 genes for which there are genetic tests used in medical practice (3). Almost daily, discoveries are announced for gene variants that affect the risks of diseases of major public health importance, ranging from adult chronic diseases such as cancer and diabetes to infectious and immunologic disorders to diseases affecting the health of infants and children. Most discoveries for gene variants are based on studies of high risk families or selected groups. In order to translate the results of this genetic research into opportunities for treating and preventing disease and promoting health, population-based epidemiologic studies are increasingly needed to quantify the impact of gene variants on the risk of disease, death and disability, and to identify and quantify the impact of modifiable risk factors that interact with gene variants. The results of such studies will help medical and public health professionals to better target medical, behavioral and environmental interventions.

Epidemiologic studies are also required in the process of clinical validation of new genetic tests, to measure the disease positive and negative predictive values of new tests, to monitor utilization of genetic tests, and to determine the safety and effectiveness of genetic tests and services in different populations (4). Given the massive amount of population-based data that will be generated over the next decades, we believe a coordinated global effort is needed to disseminate human genome epidemiologic information in order to keep up with the progress of the Human Genome Project and its accompanying gene discoveries. For this to be accomplished, there must be collaboration among epidemiologists, clinical geneticists, basic scientists, medical and public health practitioners from government, professional, academic, industry and consumer organizations worldwide.

Epidemiologists and geneticists must play key roles in collaboration with government agencies, industry, professional and consumer organizations. At the present time, much of the ongoing clinical genetic research is not population-based, and much of population-based epidemiologic research does not assess genetic risk factors. A combined genetic-epidemiologic approach is essential for a better understanding of disease etiology and for the development of molecular diagnostics. Data generated from such collaboration are urgently needed for the development of medical and public health policy. For example, issues are currently debated regarding population-based genetic testing for breast cancer in relation to BRCA1 (5), Alzheimer's disease in relation to the Apolipoprotein E-E4 allele (6), and iron overload in relation to the hemochromatosis gene (7). Given the paucity of population-based epidemiologic data regarding the frequency, disease risks and environmental interactions for many newly discovered human gene variants, we are concerned that appropriate health policy on the use of genetic tests may not be possible.

Human Genome Epidemiology

In recent years, we have seen the emergence of genetic epidemiology (8) and molecular epidemiology (9). While some view the two fields synonymously, others perceive major differences in their evolution, objectives, and focus (9). In fact, the authors of this editorial have had such conflicting opinions! However, there is one issue that the authors agree upon and strongly advocate: the critical role of population-based studies in translating findings of the Human Genome Project into applications for medicine and public health.

Here we use the term human genome epidemiology (HuGE) to denote an evolving field of inquiry that uses systematic applications of epidemiologic methods and approaches in population-based studies of the impact of human genetic variation on health and disease. We view human genome epidemiology as the intersection between molecular epidemiology and genetic epidemiology. The spectrum of topics addressed by investigators working on human genome epidemiology (Table 1) ranges from population-based epidemiologic research on gene variants to evaluation of genetic tests and services. The need to begin addressing the range of questions addressed by HuGE was recently emphasized in a one-day workshop at the Centers for Disease Control and Prevention, in which representatives from federal, state, academic, and industry organizations discussed and endorsed the vision of HuGENet™ (cf. acknowledgments).

Human Genetics Information on the World Wide Web

Given the fast pace of scientific research in medical genetics, Sikorski and Peters recently predicted that much of the up-to-date information required to practice medicine in the post Human Genome Project era most likely will reside on the Internet (10). Although no single Internet site can provide comprehensive information, an amazing depth of content already is on the World Wide Web. The available online information systems include those for gene mapping information (e.g., human genome database), gene sequence information (e.g., GenBank), gene nomenclature, gene mutation information (e.g., human gene mutation database), and comprehensive phenotype and gene-specific clinical and research information (e.g., Online Mendelian Inheritance in Man (OMIM) and GENLINE). Links to online genetic databases are available on the homepage of the CDC's Office of Genomics and Disease Prevention (11).

None of the online knowledge bases, however, provide epidemiologic population-specific data on disease risks, allele frequencies, laboratory quality, or the safety and effectiveness of genetic tests. The proposed Human Genome Epidemiology Network is a much needed effort to complement basic research and clinical sources of genomic information and to fill the widening gap between genetic research findings and medical and public health practice.

The Human Genome Epidemiology Network (HuGENet™)

It is upon this basis that the vision for the Human Genome Epidemiology Network (HuGENet™) has emerged. HuGENet™ is a logical outcome of the increasing interaction between medicine and public health as reflected in the recent CDC Strategic Plan "Translating Advances in Human Genetics into Public Health Action" (12). HuGENet™ represents the collaboration of individuals and organizations from diverse backgrounds who are committed to the development and dissemination of population-based human genome epidemiologic information.

The mission of HuGENet™ is threefold: 1) To establish an information exchange network that promotes global collaboration in the development and dissemination of peer-reviewed

epidemiologic information on human genes. 2) To develop an updated and accessible knowledge base on the World Wide Web. 3) To promote the use of this knowledge base by health care providers, researchers, industry, government, and the public for making decisions involving the use of genetic tests and services for disease prevention and health promotion.

A web site will be created at www.cdc.gov/genetics to facilitate collaboration, communication, and information exchange. The development and implementation of HuGENet™ will be carried out under the direction of a steering committee and an advisory committee in collaboration with multidisciplinary review groups worldwide. An informatics group will design and implement the database aspect of the HuGENet™ knowledge base. The web site will house the cumulative and changing information on epidemiologic aspects of human genes. Ultimately, the HuGENet™ site will contain: human genome epidemiology (HuGE) reviews that will provide concise peer-reviewed information on the epidemiologic aspects of human genes (see accompanying announcement); original human genome epidemiology research articles of the kinds described in Table 1; systematic reviews such as technology assessments, meta analyses, workshop summaries, consensus conferences; updated medical literature searches for HuGE information on specific genes; links to Internet sites and databases that provide useful population-based information on specific genes worldwide; commentaries, editorials, and opinion pieces; announcements for conferences, workshops and training opportunities in human genome epidemiology; funding opportunities in human genome epidemiology; and a forum of communication and dialogue.

Additional information on how to join the HuGENet™ collaboration can be found at www.cdc.gov/genetics. Authors interested in contributing HuGE reviews can consult the accompanying description of these reviews and submission instructions.

References

1. Ellsworth DL, Hallman DM, Boerwinkle E. Impact of the Human Genome Project on epidemiologic research. Epidemiol Rev 1997; 19:3-13.

2. Online Mendelian Inheritance in Man (OMIM). Bethesda, MD: National Center for Biotechnology Information: National Library of Medicine (www.ncbi.nlm.gov/OMIM).

3. Pagon, RA, Covington M, Tarczy-Hornoch P: GeneTests (formerly Helix: a Directory of Medical Genetics Laboratories, companion site of GeneClinics: Medical Genetics Knowledge Base).

4. Khoury MJ. Genetic epidemiology and the future of disease prevention and public health. Epidemiol Rev 1997;19:175-180.

5. Burke W, Daly M, Garber J, et al. Recommendations for follow up care of individuals with an inherited predisposition to cancer. II, BRCA1 and BRCA2. JAMA 1997;277:997-1003.

6. American College of Medical Genetics/American Society of Human Genetics Working Group. Statement on use of apolipoprotein E testing for Alzheimer disease. JAMA. 1995;274:1627-1629.

7. McDonnell SM, Witte D. Hereditary hemochromatosis. Preventing chronic effects of this underdiagnosed disorder. Postgrad Med. 1997;102:83-85.

8. Khoury MJ, Beaty TH, Cohen BH. Fundamentals of Genetic Epidemiology. New York, Oxford University Press, 1993.

9. Dorman JS, on behalf of the International Molecular Epidemiology Task Force. Molecular epidemiology and DNA technology transfer: A program for developing countries. In: Dorman JS (ed). Standardization of Epidemiologic Studies of Host Susceptibility. Plenum Press, 1994, pp 241-256.

10. Sikorski R, Peters R. Genomic medicine. Internet resources for medical genetics. JAMA 1997;278:1212-1213.

11. Office of Genomics and Disease Prevention homepage, Centers for Disease Control and Prevention.

12. Task Force on Genetics in Disease Prevention: Translating Advances in Human Genetics into Public Health Action: A Strategic Plan. Centers for Disease Control and Prevention 1997.

Table 1
Functions of Human Genome Epidemiology

Assess the prevalence of gene variants in different populations.

- Assess the magnitude of disease risk associated with gene variants in different populations (relative and absolute risks).
- Assess the contribution of gene variants to the occurrence of the disease in different populations (attributable risks).
- Assess the magnitude of disease risk associated with gene-gene and gene-environment interaction in different populations.
- Assess the validity of genetic tests in different populations (disease positive and negative predictive values).
- Evaluate the magnitude and determinants of the utilization of genetic tests and services in different populations.
- Evaluate the impact of genetic tests and services on morbidity, disability, mortality and cost in different populations.

Acknowledgments

We appreciate the participation and comments of the following individuals who attended the CDC HuGENet™ workshop on November 20, 1997: Larry Gelbert Ph.D., , Bristol-Myers Squibbs; Ronald LaPorte, Ph.D., University of Pittsburgh, Graduate School of Public Health, Department of Epidemiology; Eugene Lengerich, D.V.M, M.S., North Carolina Department of Health and Human Services, Michele Puryear M.D., Ph.D., Maternal and Child Health Bureau, Health Resources and Services Administration, Daniel W. Nebert, M.D., Center for Environmental Genetics, and Departments of Environmental Health and of Pediatrics,University of Cincinnati; Iris Obrams, M.D., Ph.D, and Susan Nayfield M.D., National Cancer Institute; Roberta A. Pagon, M.D., University of Washington School of Medicine, Division of Medical Genetics; Robert J. Robbins, Ph.D., Vice President for Information Technology, Fred Hutchinson Cancer Center; Allen Wilcox, M.D., Ph.D., National Institute of Environmental Health Sciences

Announcement

The American Journal of Epidemiology and Epidemiologic Reviews, in collaboration with the Centers for Disease Control and Prevention announce the launching of the Human Genome Epidemiology Network (HuGENet™). The rationale, purpose and goals of HuGENet™ are outlined in the accompanying editorial. We invite authors to submit human genome epidemiology (HuGE) reviews for possible publication in the American Journal of Epidemiology or Epidemiologic Reviews. These submissions will be peer-reviewed, and if accepted, will be simultaneously published in the HuGENet™ knowledge base on the World Wide Web at www.cdc.gov/genetics. The suggested components for a HuGE review are outlined in the accompanying table. The topics of these reviews can range from genetic variants or mutations associated with high disease risks (e.g., cystic fibrosis) to polymorphisms and normal variants (such as Apolipoprotein E-E4) associated with variable disease risks. In the case of genes associated with multiple disease outcomes (e.g., BRCA1 and breast cancer and ovarian cancer), HuGE review authors may elect to restrict their reviews to one or more disease entities as long as they acknowledge the association with other diseases. In addition, when multiple gene variants are associated with different diseases, reviewers may focus on one or more of these variants in relation to one or more diseases (e.g., Apolipoprotein E-E4 allele and Alzheimer's disease). Authors should acknowledge the presence of other gene variants (e.g., Apoliprotein E-E3 or E2) and the association with other diseases (e.g., coronary heart disease).

The objective of a HuGE review is to identify a gene's allelic variants and describe what is known about the frequency of these variants in different populations, to identify diseases that these variants are associated with and to summarize the magnitude of risks and associated risk factors. A crucial component of a HuGE review is to identify gaps in the epidemiologic knowledge base for human gene variants in order to stimulate further work in these areas. We expect these reviews to be concise (10-15 double spaced pages). While no formal meta analytic methods are required, authors are expected to describe the sources of information they searched, their criteria for including some papers and excluding others, and their criteria for evaluating the quality of publications and the methods they used to summarize data and draw conclusions. Tables and figures are welcome in summarizing and evaluating the quality of epidemiologic information. We expect authors to provide up-to-date references (fewer than 50). We also expect relevant links to World Wide Web resources, including genetics databases, online resources, educational materials, policy statements, and support groups.

If you are interested in submitting a HuGE review, please send an electronic message of your intent to HuGENet™ coordinator at genetics@cdc.gov to sign up for a HuGE review. A HuGE review is due within 6 months. An updated list of HuGE reviews that are under development will be kept on the HuGENet™ web site at www.cdc.gov/genetics. This will be crucial to keep track of HuGE reviews under development and to avoid potential duplication and overlap. Please submit completed HuGE reviews to editor, American Journal of Epidemiology following the usual guidelines for submission. Accepted papers will be simultaneously published in either the American Journal of Epidemiology or Epidemiologic Reviews and HuGENet™.

Table 2

Suggested Format for a HuGE review

1. ***At-A-Glance***
One page synopsis of the issues discussed in items 2-10 below with a brief statement on each of these items

2. ***Gene***
Identify the gene being reviewed and provide a brief review of its chromosome location, the gene product and its function, if known.

3. ***Gene variants***
List known allelic variants and summarize known information on the frequency of homozygosity and heterozygosity of these variants in different populations and ethnic groups.

4. ***Disease(s)***
Identify the disease(s), disorders this gene is associated with. Briefly summarize the descriptive epidemiology and confirmed and suspected risk factors (including other genes).

5. ***Associations***
Summarize the magnitude of the association between the allelic variants and various diseases-absolute, relative and attributable risks in different populations. Comment on the quality and methodology of studies.

6. ***Interactions***
Discuss whether the allelic variants of this interact with any of the disease known risk factors including other genes and environmental factors. Summarize the magnitude of such interactions.

7. ***Laboratory tests***
Summarize the sensitivity, specificity, and predictive values of different tests available for this gene, including biochemical, molecular, and other tests in different populations.

8. ***Population testing***
Summarize population-specific data on the magnitude and determinants of testing for allelic variants of this gene and the impact of testing on public health (morbidity, mortality, disability), including policy statements, recommendations, and legislation (including mention of available interventions).

9. ***References***
Include relevant Medline citations (fewer than 50).

10. ***Internet sites***
Include relevant links to various genetics databases, online resources, educational materials, policy statements, and support groups.